

# White Paper Report

Report ID: 104045

Application Number: HW5000909

Project Director: Thomas Elliott (tom.elliott@nyu.edu)

Institution: New York University

Reporting Period: 5/1/2009-4/30/2010

Report Due: 7/31/2010

Date Submitted: 8/16/2011

# Final Performance Report

Report ID: 94247

Application Number: HW5000909

Project Director: Thomas Elliott (tom.elliott@nyu.edu)

Institution: New York University

Reporting Period: 5/1/2009-4/30/2010

Report Due: 7/31/2010

Date Submitted: 7/7/2010

Final Performance Report

Grant number: HW-50009-09

Title of Project: Epigraphic Interoperability Workshops

Name of Project Directors: Dr. Thomas Robert Elliott (NYU); Prof. Dr. Christian Witschel (Heidelberg)

Name of Grantee Institution: New York University

Date Report is Submitted: 7 July 2010

# Introduction

This project comprised a pair of workshops involving representatives of the Pleiades Project, the *Epigraphische Datenbank Heidelberg* (EDH) and an outside expert representing the EpiDoc community. EDH is one of the largest and most sophisticated scholarly databases of ancient Roman texts originally inscribed or incised on durable materials like stone and metal (60,000 inscriptions and growing). The international Electronic Archive of Greek and Latin Epigraphy (EAGLE), launched by the *Association Internationale d'Épigraphie Grecque et Latine* (AIEGL) in 2003, has relied upon EDH from the beginning to serve as the primary repository for Latin and bilingual (Latin-Greek) inscriptions associated with all the provinces of the Roman Empire. Pleiades is a multi-institutional project (headquartered at New York University's Institute for the Study of the Ancient World, ISAW) that gives scholars, students, teachers, school pupils and enthusiasts worldwide the ability to use, create and share historical geographic information about the Greek and Roman World. EpiDoc represents a growing, global collaboration of humanists and information technologists whose joint aim is the creation of flexible but rigorous standards and tools for the digital encoding and interchange of scholarly and educational editions of ancient texts, especially those inscribed on durable materials or written on papyrus.<sup>[1](#)</sup>

Our goal for this workshop series was to take the first planning steps on the road to open, seamless cross-project interoperability between EDH, Pleiades and appropriately licensed EpiDoc-conformant publications.

## Project activities and accomplishments

The project team completed all of the activities outlined in the proposal work plan; i.e., two workshops with associated pre- and post-coordination and discussion.

The first workshop was held in [Heidelberg](#), 14-17 July 2009, as planned. Tom Elliott and Sean Gillies represented ISAW. Participating Heidelberg staff included Christian Witschel, James Cowey, Francisca Feraudi-Gruénais, Brigitte Gräf and Andreas Dafferner. We were also joined by Gabriel Bodard from King's College, London, representing the EpiDoc community. The Heidelberg workshop gave rise to list of specific technical issues to be explored and associated tasks to be accomplished during the interval between workshops using non-grant funds as resources permitted. Major areas of focus included: procedures for ingest into EDH of texts from the EpiDoc-conformant *Inscriptions of Roman Tripolitania 2009*, and procedures for computationally assisted linking to Pleiades of EDH records for inscriptions from the province of Moesia Inferior.<sup>[2](#)</sup>

The second workshop was held in [New York](#), 4-6 March 2010. This date was moved back from the original plan of January/February 2010 because of a change of IT personnel in Heidelberg. The long-serving IT specialist Andreas Dafferner announced his intention in November 2010 to leave his post for a different job in Heidelberg, effective January 2010. This development rendered

impossible progress on geographic aspects of the agenda agreed at the first workshop. All parties agreed that a delay to allow Heidelberg to recruit a replacement would be beneficial to the project, and at the beginning of February 2010, Frank Grieshaber took up the appointment. Thanks to the delay, Grieshaber was able to join Feraudi-Gruénais and Cowey in traveling to New York, where they met with Gillies and Elliott. This meeting was also productive, resulting in a series of technical decisions that lay the groundwork for future collaboration and set the stage for a forthcoming effort to achieve full interoperability between the EDH and Pleiades web applications.

Outlined below are the major areas of collaboration and progress that have emerged from this series of workshops.

### ***Geography and EDH/Pleiades interoperation***

The original plans outlined in the application have been adhered to.

On the basis of discussions and decisions reached during the workshops a new geographical database has been created with non-grant funds and is available in beta version which is being tested in Heidelberg. This will change some of the data entry practice in EDH in as much as each and every combination of place names (ancient findspot, modern findspot and *Fundstelle*) will be entered only once and receive a unique identity number within the geographical database. Pleiades identifiers, as well as site coordinates drawn from a range of appropriate resources, will also be recorded in this database thereby ensuring both control over the geographic content associated with EDH records and interoperation with Pleiades and other emerging sources of open geographic information on the web. The EDH geographic identity number will be used in the core epigraphic database wherever the combination of geographical data is applicable, thus minimising orthographic variation (heretofore, only placename strings were recorded). In essence the geographic database will establish an ever-growing authority list for the place names associated with the epigraphy in their correct combination with one another. Future site enhancements will exploit this infrastructure to facilitate dynamic mapping and crosslinks to Pleiades resources associated with individual inscriptions (e.g., the ancient places associated with the authoring or discovery of the texts).

Concurrent with the collaborative effort to define a new geographical database for the EDH content, both parties also collaborated to devise a Pleiades interface to facilitate supervised matching of Pleiades geographic resources with placename input. Specifications for this interface were worked out during the workshops. Subsequently, using non-grant funds, the Pleiades team developed a simple web form that takes as input modern and/or ancient names, as well as an optional modern country code. The form returns a query result including URIs for all place resources in Pleiades with a matching ancient name and URIs for all modern-name+country matches from the open-content GeoNames.org service. The user can work with these results in multiple formats, including text, JavaScript Object Notation (JSON; suitable for scripting into custom mapping applications) and Keyhole Markup Language (KML;

suitable for visualization in Google Earth). This "place-matching interface" is now being tested by the EDH/Pleiades team, and by other members of the Pleiades community at <http://pleiades.stoa.org/@@place-match-form>. The ISAW team expects to expand its capabilities in coming months as part of its recently-awarded NEH Preservation and Access grant for "Pleiades 2" (PW-50557-10). Refinements to the interface will be solicited and vetted in cooperation with EDH and other individuals and organizations with existing data and a concrete use case for Pleiades interoperability, including the Harvard-based *Digital Atlas of Roman and Medieval Civilization* project.<sup>3</sup>

Significant discussion also revolved around mechanisms to facilitate web-based citation (linking) across the two sites, as well as means of surfacing the scholarly function of these links for the benefit of users (e.g., assertions about findspot relationships or the attestation of a placename). Accordingly, EDH has implemented simple web addresses for each inscription that hide the mechanics of query lookups and software implementations (e.g., <http://edh-www.adw.uni-heidelberg.de/EDH/inschrift/009931>). This step, which follows the recommendations of the World-Wide Web Consortium, brings EDH into alignment with Pleiades' commitment to the kind of stable, web-actionable identifiers essential to online scholarly citation. Future enhancements to EDH, resources permitting, will include development of another feature pioneered by Pleiades: Atom-format web feeds summarizing individual resources and the relationships the project asserts between them and other web-based resources. This approach was prototyped during the Concordia Project, jointly funded by the NEH and the UK JISC. Its implementation will facilitate automated harvesting of EDH content and relationships for presentation to Pleiades users in the context of individual resources. It will also provide web-based search engines, like the GoogleBot, with a comprehensive index of EDH content, thereby enhancing discoverability on the open web.<sup>4</sup>

### ***EDH/EpiDoc interoperation***

The original plans outlined in the application have been adhered to.

Prior to and during the July 2009 workshop, Gabriel Bodard produced XML Stylesheet Language Transformations (XSLT) to accomplish conversion of 1) IRT 2009 EpiDoc XML to the database structure of EDH metadata, and 2) txt files containing the relevant Latin and/or Greek content of the specific inscription. This code effectively documents the piecewise relationships between the two formats. In the course of November and December of 2009 this XSLT was refined and improved by cooperation between Gabriel Bodard and James Cowey. The result produced exports which were then checked and cleaned up by Francisca Feraudi-Gruénais and Brigitte Graef for import into EDH (for concrete results and links see below). This cleaning process has been documented and can thus contribute to further improvements to the XSLT in order to minimise additional clean up in Heidelberg for the needs of EDH.

Inevitably XSLT for transforming the IRT 2009 EpiDoc XML to EDH formats will not be exactly the same as XSLT for transforming other EpiDoc resources to EDH format, but the tweaks required will be minimised because the EDH

requirements remain the same and the experience gathered in crosswalking the IRT 2009 data leads to improvements in the transformation process. This allows us to look at future cooperation projects in the knowledge that data export and import can be implemented with increasingly less time and effort involved in preparing and carrying out such tasks.

Development of detailed specifications for export of EpiDoc from EDH were deferred for two reasons. It seemed prudent to wait for the release (scheduled 2010) of "round-trip" EpiDoc conversion software currently under development by a multi-institutional project called "Integrating Digital Papyrology" (IDP), which is led by Duke University with funding from the Andrew W. Mellon Foundation. Moreover, a meeting scheduled for Rome in November 2010 presents an opportunity to develop strategy for an EpiDoc export solution (as well as cross-resource search and map visualization) that would meet the needs of the entire EAGLE consortium of epigraphic databases, rather than just EDH (see further, below).<sup>5</sup>

## **Audience, Continuation and Long-term Impact**

Pleiades, EDH and the EpiDoc community serve overlapping, international audiences of scholars, students and enthusiasts. Although it is not possible to provide quantitative documentation of concrete effects on specific audience demographics for a planning exercise of this nature, we are confident that we have laid the groundwork for significant impact in future. Just as epigraphy provides indispensable evidence for the study of ancient geography, so ancient geography is an essential aid in the study of epigraphic material. Consequently, future interoperational modes that exploit links between resources and provide for the free flow of information from one application to another seem likely not only to produce positive value for our audiences, but also to serve as salutary examples for other online scholarly reference works. Key groups outside the immediate project team support this assessment.

Each project funded by the Heidelberg Academy of the Sciences is monitored by an external review board, or "Kommission", which meets yearly to discuss progress, developments, planning and the future direction of the project. This year, the EDH Kommission met in February and was briefed on the cooperation between EDH and Pleiades. This aspect of EDH's work was praised for its potential, from the point of view of both the project's internal goals and the wider needs of the EAGLE consortium in which it participates.

In the space of time between the two workshops, the EDH team has had meetings with Prof. Mark Depauw from Leuven. He is in charge of the Trismegistos project, whose mission includes the task of providing an essential prerequisite for linking external databases with one another: assignment of unique identifiers to text editions of papyri, ostraca and related text-bearing materials (bones, wood etc.). He has started a pilot project with various regions in EDH (Macedonia, Histria and Venetia). This involves the export of a data set from EDH which is then checked for any irregularities before being returned with a TM number (Trismegistos number) for each edition. This Trismegistos number is then recorded in the EDH database, which has been

modified to accommodate this new field. The establishment of such identifiers - already successfully exploited by IDP to link three seminal digital resources in papyrology -- sets the stage for high-quality, federated search across multiple databases with full disambiguation of doublets and related entries.[6](#)

Together, the outcome of the meetings supported by this grant, and the separate collaboration with Trismegistos, led Silvia Orlandi (Professor of Latin Epigraphy at La Sapienza in Rome) to include the inter-related issues of interoperation, federated search, shared geographic visualization and EpiDoc import/export on the agenda for the next meeting of the EAGLE consortium, scheduled for November 2010. Elliott, Witschel, Feraudi-Gruénais, Grieshaber, Cowey and other team members will participate.

In sum, the potential for building on this workshop series is significant and growing.



## Notes

<sup>1</sup> EDH: <http://www.uni-heidelberg.de/institute/sonst/adw/edh/indexe.html> ;  
EAGLE: [http://eagle.let.uniroma1.it/Italiano/index\\_it.htm](http://eagle.let.uniroma1.it/Italiano/index_it.htm); Pleiades:  
<http://pleiades.stoa.org> ; EpiDoc: <http://epidoc.sf.net>

<sup>2</sup> IRT 2009: *Inscriptions of Roman Tripolitania*, by J. M. Reynolds and J. B. Ward-Perkins, enhanced electronic reissue by Gabriel Bodard and Charlotte Roueché (2009). ISBN 978-1-897747-23-3. Available:  
<http://irt.kcl.ac.uk/irt2009/>

<sup>3</sup> DARMC: <http://darmc.harvard.edu/icb/icb.do> (the DARMC/Pleiades collaboration is funded as part of Pleiades 2)

<sup>4</sup> W3C "Cool URI" recommendations: <http://www.w3.org/Provider/Style/URI>;  
Atom: <http://tools.ietf.org/html/rfc4287> ; more on this aspect of the Concordia project: <http://www.atlantides.org/trac/concordia/wiki/ConcordiaAtomFeeds>

<sup>5</sup> IDP: <http://idp.atlantides.org>

<sup>6</sup> Trismegistos: <http://www.trismegistos.org/>

<sup>7</sup> HD059004:  
<http://edh-www.adw.uni-heidelberg.de/offen/suchen2.html?hdnr=059004> or  
<http://edh-www.adw.uni-heidelberg.de/EDH/inschrift/059004>

<sup>8</sup> HD059063:  
<http://edh-www.adw.uni-heidelberg.de/offen/suchen2.html?hdnr=059063> or  
<http://edh-www.adw.uni-heidelberg.de/EDH/inschrift/059063>

## **APPENDICES**

### ***Samples of completed work***

Transformation and export of text and metadata of [IRT](#) and import of said material into [EDH](#). At present in EDH the numbers [HD0590047](#) thru [HD0590638](#) demonstrate examples of completed integration of IRT texts into EDH.

Agenda, notes and products from the two workshops -- necessarily rough and ready -- are publicly accessible via links on the following web page:  
[http://www.atlantides.org/trac/concordia/wiki/NEH\\_DFG](http://www.atlantides.org/trac/concordia/wiki/NEH_DFG).